Hicks, D., & Wang, J. (2009). Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project. Georgia: Georgia Institute of Technology.

Meester, W. (2013). Value of Bibliometrics.Towards a comprehensive citation index for the Arts & Humanities. *Research Trends,* (32). Retrieved from http://www.researchtrends.com/issue-32-march-2013/towards-a-comprehensive-citation-index-for-the-arts-humanities/ website:

Moed, H. F., Linmans, J., Nederhof, A., Zuccala, A., López Illescas, C., & de Moya Anegón, F. (2009). Options for a Comprehensive Database of Research Outputs in Social Sciences and Humanities. Research report to the Project Board of the Scoping Study "Towards a Bibliometric Database for the Social Sciences and the Humanities" set up by the Standing Committees for the Social Sciences and the Humanities of the European Science Foundation (ESF). Version 6 April 2009. Leiden, Granada: Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands; SCIMago Research Group, CSIC Madrid and University of Granada, Spain.

Rey, O. (2006). What is 'good' research in education? *Veille scientifique et technologique.* Lyon: Institut National de Recherche Pédagogique.

Yousfi, S., Koopmann, B., & Amelang, M. (2005). *Correlates of item validity. On the eminent importance of global self-ratings.* University of Standford; University of Heidelberg. Heidelberg, Stanford.

A Prototype Empirical Framework of Intrinsic and Extrinsic EERQI Indicators

Ton Mooij

Short Summary

The research question to be answered in the context of the EERQI project is: What do statistical analyses show us about the relationships between intrinsic and extrinsic indicators of quality and what does this mean when constructing a prototype EERQI framework?

The pilot study involved the scoring on both intrinsic and extrinsic indicators for 177 research documents or articles written by 268 authors. Intrinsic data were gathered by peer review and extrinsic data were collected from the Internet. Confirmatory Factor Analysis (CFA) resulted in a measurement model containing three intrinsic and two extrinsic latent factors. There are significant correlations between the intrinsic latent factors and between the extrinsic latent factors; however, no significant correlations have been found between intrinsic and extrinsic factors. This outcome underlines the notion that intrinsic indicators may add specific quality information to an EERQI prototype that consists solely of extrinsic indicators, and vice versa. Testing by means of a structural model revealed that the more a reviewed document is related to the reviewer's own area of research, the higher the score the reviewer gives the document with respect to 1) significance, originality and consistency and 2) methodological adequacy. No relationships were found between the reviewer's own area of research and the extrinsic latent factors. These effects on the two intrinsic latent factors indicate that there may be some subjective evaluation bias in peer reviewing.

The conclusion is that the outcomes of the statistical analyses seem plausible and support the validity of the conceptual framework. An initial prototype EERQI framework has been constructed, which is in line with the main goal of the EERQI project. Although the pilot had some methodological limitations, the present empirical outcomes are promising for future EERQI developmental and research activities, which could, for example, also integrate semantic latent factors and indicators.

1 Introduction

Impact indicators in educational research such as those based on data from Social Science Citation Index (SSCI) were long based on measures of citations to documents in specific scientific journals. Nowadays, search engines automatically use various types of 'objective', 'external', or 'extrinsic characteristics', e.g. the bibliometric or semantic characteristics of publications or documents found in many different Web-based sources. The goal of the international 'European Educational Research Quality Indicators' project (EERQI; FP7 # 217549) is to improve citation-only assessments of the quality or impact of educational and other research (cf. Gogolin, 2008; Gradmann, Sieber, & Stoye, 2011). Therefore, in addition to bibliometric and citation indicators, efforts were also made to distinguish specific indicators reflecting the more 'subjective', 'internal', or 'intrinsic' quality of research documents. Intrinsic indicators refer to the content of a publication or research document and are supposed to explicate or describe such aspects as rigour, originality, significance, integrity, and style (cf. Bridges, 2009).

In earlier EERQI papers (Mooij, 2008a, 2008b), I outlined a possible approach to empirically exploring and analysing relationships between sets of intrinsic and extrinsic indicators of the quality of research documents. I also analysed the statistical relationships between both intrinsic and extrinsic indicators in order to construct an initial EERQI prototype framework (Mooij, 2011). In this chapter I concentrate on the main aspects and outcomes of the empirical research involved in constructing this prototype. The research question to be answered is: What do statistical analyses show us about the relationships between intrinsic and extrinsic indicators of quality and what does this mean when constructing a prototype EERQI framework?

2 Intrinsic and extrinsic quality indicators

2.1. Intrinsic indicators

In the EERQI project, intrinsic indicators were chosen to operationalise the concepts: methodology, results, discussion, originality, significance, validity, and miscellaneous. Twenty items were devised to assess these seven quality concepts: see for these concepts and items Table 1. The answer alternatives for each item were: 'not relevant for this text' (=0), 'very poor' (=1), (2), (3), 'average' (=4), (5), (6), and 'excellent' (=7). A final item, item 21, allowed peer reviewers to indicate how closely the document they had evaluated related to their own area of research. Here the answer categories were: 'Very closely' (=1), 'Closely' (=2), 'Less closely' (=3), 'Not at all' (=4). The complete dataset resulting from the final pilot in the EERQI project consists of 177 research documents or articles written by a total of 268 authors. Peer reviewers scored these documents with respect to all 21 items. For each document, peer review evaluation scores were aggregated by calculating their mean across reviewers.[26]

---

[26] The dataset containing both the intrinsic and extrinsic scores of 177 research documents became available on 2 March 2011. The dataset contains scores by peer reviewers who are partners in the EERQI project or attended the European Conference on Educational Research in 2010. Some of the reviewers scored two or more research articles. If available per document, the scores of various reviewers were aggregated. It seems that value 0 ('not relevant for this text') was included in these scores, however. This problem could not be avoided because only the aggregated data were available. The 177 documents represent three different European languages. In combination with the small number of reviewers, the actual data structure does not permit assessment of interobserver reliability or multilevel analyses between and within languages and/or reviewers, respectively.

Table 1 – Concepts and items assessing intrinsic quality (n documents=171)

Concept_var.	Description of variable or item	Min.	Max.	M	SD
1 Methods_1	The methods are intelligibly described	.00	7.00	4.02	2.03
2 Methods_2	The method / approach is appropriate	.00	7.00	4.70	1.63
3 Methods_3	The method / approach is accurate	.00	7.00	4.34	1.78
4 Results_1	The results are completely described	.00	7.00	4.51	1.66
5 Results_2	The results are correctly described	.00	7.00	4.53	1.67
6 Discussion_1	The study's method is reflected in an appropriate way	.00	7.00	3.94	1.82
7 Discussion_2	The study's results are reflected in an appropriate way	.00	7.00	4.51	1.69
8 Discussion_3	The pattern of reasoning is consistent	1.00	7.00	5.48	1.10
9 Discussion_4	The discussion shows a critical evaluation of the work	.00	7.00	4.67	1.47
10 Originality_1	The study shows new approaches in its methodological procedures	.00	7.00	3.39	1.63
11 Originality_2	The study shows new approaches in the structure of its argumentation	.00	7.00	4.16	1.35
12 Originality_3	The study contributes innovative ideas for the state-of-art in its research area	.50	7.00	4.52	1.33
13 Significance_1	The study contributes to the development of its research field	1.00	7.00	5.02	1.28
14 Significance_2	The study makes a significant contribution to the latest discussions within the research field	1.00	7.00	4.82	1.30
15 Significance_3	The study makes a significant contribution to the latest discussions within the educational policy field	.00	7.00	4.62	1.52
16 Significance_4	The study makes a significant contribution to the latest discussions within the educational practice field	.00	7.00	4.51	1.57
17 Validity_1	How do you evaluate the article concerning its Rigour?	.00	7.00	4.72	1.38
18 Validity_2	How do you evaluate the article concerning its Originality?	1.00	7.00	4.82	1.05
19 Validity_3	How do you evaluate the article concerning its Significance?	1.00	7.00	5.03	1.22
20 Miscellaneous2	Comparing this article to an article representing good research, where would you place it on a scale from 1 to 7, with 7 being excellent quality and 1 being bad quality?	1.00	7.00	4.61	1.11
21 Miscellaneous1	The reviewed article is related to my own area of research…	1.00	4.00	2.40	0.54

In the univariate analysis using the Statistical Package for the Social Sciences (SPSS, version 17.0), only documents without system-missing values were used, which resulted in item-specific information for 171 documents. Table 1 also presents the descriptive statistics of these intrinsic items. The means vary from around 4 (average) to 5; standard deviations vary from 1.05 to 2.03.

2.2. Extrinsic indicators

Extrinsic indicators usually measure aspects of research documents such as number or distribution of citations (per author; across authors; per document; hits resulting from search engines for a paper or author/combination of authors, and so forth). The information on extrinsic indicators was provided per author. Because research documents constitute the unit of analysis, the extrinsic information was aggregated per document. When there was more than one author per document, the available information per indicator was aggregated by totalling the scores of the authors per document.[27] The dataset of 2 March 2011 contains information about 12 extrinsic indicators. Five of these were neglected.[28] Information about the remaining seven extrinsic indicators, their range of scores, means and standard deviations is given in Table 2.

Table 2 – Variables assessing extrinsic quality (n documents=171)

Variable name	Description	Min.	Max.	M	SD
1 Cit/paper	Citations per paper without self-citations using full title of the article	.00	804.81	18.48	64.36
2 WebMennAuth	Web mentions of author in search engine BING; number of URLs of pages matching the query submitted	2.00	1791.00	352.23	280.33
3 WebMentTitle	Web mentions of article title in search engine BING; number of URLs of pages matching the query submitted.	.00	1046.00	25.24	131.59
4 GoogleHits	Google Web Search results	.00	3265.00	219.91	448.85
5 MetagerHits	Metager hits	.00	133.00	4.74	16.16
6 CiteULikeHits	Mentions of article CiteULike	.00	486.00	21.32	60.55
7 LibraryThingHits	Mentions of article LibraryThing	.00	651.00	29.34	89.95

The variable 'number of citations per paper' [Cit/paper] has a very skew distribution to the right.[29] The respective scores were therefore transformed by taking their square roots. The range of the transformed scores is 0.00 – 28.37 with Mean 3.24 and SD 2.83. Principal factor analysis was used to explore the relationships between the seven extrinsic variables listed in Table 2. The variables WebMentTitle and MetagerHits are not related to the other variables or only to a very limited extent. Given the present focus, it was decided to drop these two variables.

The Eigenvalues and percentages of variance of the remaining five variables point to the presence of two underlying factors: see Table 3.

Table 3 – Eigenvalues and % of variance for extracted factors of five extrinsic variables

Factor	Eigenvalue	% of Variance	Cumulative %
1	2.612	52.236	52.236
2	1.152	23.039	75.276
3	.606	12.126	87.401
4	.444	8.882	96.283
5	.186	3.717	100.00

The loadings of the five variables on the two factors were rotated (oblique, geomin) within the EFA procedure of statistical program MPlus 6.1: see Table 4. The results in Table 4 illustrate that 'Citations per paper (without self-citations)' and 'Web mentions of author in search engine BING' represent factor 1, whereas the second factor represents numbers of hits by three other search engines.

Table 4 – Factor loadings of extrinsic variables after oblique (geomin) rotation

| Variable name	Description	Factor 1	Factor 2
Cit/paper (sqrt)	Citations per paper without self-citations using the full title of the article	0.921	-0.001
WebMennAuth	Web mentions of author in search engine BING; number of URL's of pages matching the query submitted	0.405	0.098
GoogleHits	Google Web Search results	0.023	0.947
CiteULikeHits	Mentions of article CiteULike	0.000	0.689
LibraryThingHits	Mentions of article LibraryThing	-0.112	0.867

[29] The value '0' may reflect 'missing value' or 'no hits'/'no citations'. In this paper, the latter ('no #') is assumed.

3 Relationships between intrinsic and extrinsic indicators

3.1. Modelling intrinsic and extrinsic latent factors

The study used the results of an earlier factor analysis based on the intrinsic variables of Table 1 to construct a measurement model with two intrinsic factors and two extrinsic factors (cf. Table 4). The model is given in Figure 1.

Figure 1 – Graphic presentation of CFA measurement model (4 latent factors)
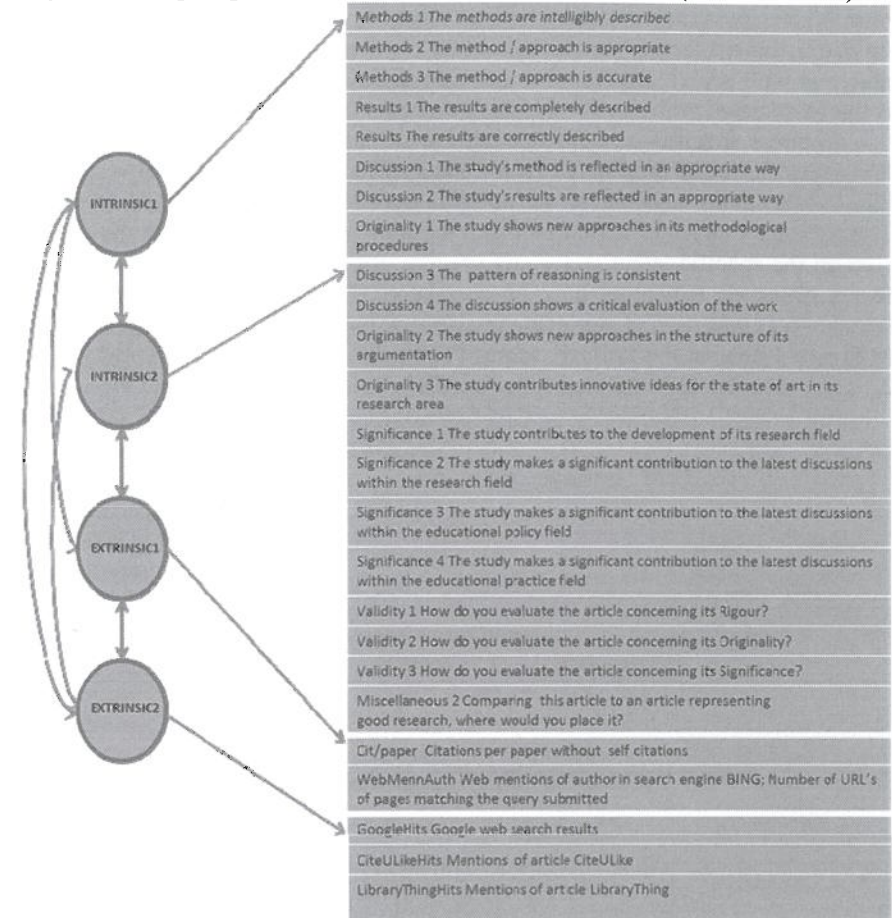
In Figure 1, latent factor 'Intrinsic1' represents the intrinsic indicators methodological adequacy, completeness and correctness of reporting results, appropriateness of discussion, and originality with respect to methodological procedures. *Intrinsic1* therefore indicates *methodological adequacy of the document*. Latent factor 'Intrinsic2' stands for logical consistency, critical evaluation, innovation, various types of significance and overall evaluation of the information in a document. *Intrinsic2* thus represents *significance, originality and consistency of the document*. Furthermore, latent factor 'Extrinsic1' refers to number of citations per document without self-citations and Web mentions of author by search engine BING. *Extrinsic1* therefore indicates *number of citations and Web mentions by BING*. Latent factor 'Extrinsic2' rather univocally represents number of hits obtained with search engines Google, CiteULike, and LibraryThing. *Extrinsic2* is thus associated with *number of hits in three specific search engines*.

Figure 1 specifies a 'Confirmatory Factor Analysis' (CFA) to check the relationships between each latent factor and specific indicators or observed variables, while taking account of the correlations between various latent factors.[30] The variance of each observed indicator variable is explained by both the regression on the specific latent factor and specific error variance; error variances between observed indicators may be correlated. The statistical program MPlus (version 6.1) was used to simultaneously check the fit of the measurement model in Figure 1 against the intrinsic scores (Table 1) and the extrinsic scores (Table 2). The outcomes of Maximum Likelihood analysis are given in Table 5.

Table 5 – ML parameter estimates (standardised) of measurement model Fig. 1

Latent factors Indicators	INTRINS1: Methodolog. adequacy	INTRINS 2: Sign./orig./ consist.	EXTRINS1: # cit./Web BING	EXTRINS2: Hits 3 searc. eng.	R^2
Methods_1	0.912**				0.832**
Methods_2	0.826**				0.683**
Methods_3	0.882**				0.777**
Results_1	0.784**				0.615**
Results_2	0.791**				0.626**
Discussion_1	0.881**				0.777**
Discussion_2	0.781**				0.609**
Discussion_3		0.656**			0.430**
Discussion_4		0.612**			0.375**
Originality_1	0.776**				0.603**
Originality_2		0.796**			0.634**
Originality_3		0.873**			0.763**
Significance_1		0.900**			0.809**
Significance_2		0.910**			0.829**
Significance_3		0.809**			0.654**
Significance_4		0.721**			0.520**
Validity_1		0.542**			0.294**
Validity_2		0.785**			0.616**
Validity_3		0.842**			0.708**
Miscellanous2		0.840**			0.706**
Cit/paper (sqrt)			0.592**		0.350**
WebMennAuth			0.685**		0.469**
GoogleHits				0.980**	0.960**
CiteULikeHits				0.674**	0.455**
LibrarThingHits				0.803**	0.645**

Factor covariances (correlations)

	INTRINSIC1	INTRINSIC2	EXTRINSIC1
INTRINSIC2	0.631**		
EXTRINSIC1	0.239*	0.148	
EXTRINSIC2	0.147	0.085	0.460**

Fit indices: X^2 (269)=1028.656 (p=0.000); RMSEA=0.129; SRMR=0.072.
* $0.01 \leq p \leq 0.05$; ** $p < 0.01$.

The overall fit of the model is reflected in two statistical indices, the 'Root Mean Square Error of Approximation' (RMSEA) and the 'Standardized Root Mean Square Residual' (SRMR): see the note following Table 5. Both measures are

[30] In the measurement model of Figure 1, the relationships between the four latent factors are standardised to facilitate their interpretation as correlations. Correlations between factors are free to vary. These correlations are represented by the two-way arrows between all pairs of latent factors. The regressions of each of the indicator variables on their respective latent factor are represented by one-way arrows. The total variance of each factor is set to 1.

related to the Chi-Square statistic. Both indices are influenced by the sample size, which implies that a smaller sample results in a less favourable fit. Generally, a value above 0.10 on both indices is considered to indicate a bad fit. With respect to the results in Table 5, it can be seen that RMSEA=0.129 and SRMR=0.072.

Table 5 furthermore demonstrates a strong correlation between the two intrinsic factors (0.631) and a weaker correlation between the two extrinsic factors (0.460). The correlation between Intrinsic1 (methodological adequacy of the document) and Extrinsic1 (number of citations and Web mentions by BING) is also significant (0.239; p≤.05). This outcome illustrates some overlap between intrinsic and extrinsic indicators, a finding that merits more attention for reasons of both EERQI interpretation and modelling.

The other correlations between intrinsic and extrinsic latent factors are not significant statistically. This implies that the use of intrinsic indicators may add quality information to an EERQI that consists solely of extrinsic indicators, or that the introduction of extrinsic indicators may add quality information to an EERQI containing only intrinsic indicators.

Given the data available and the small sample size, the overall results in Table 5 confirm the first empirical check of the validity of the measurement model in Figure 1. Moreover, the confirmatory factor loadings and the variances explained per indicator (R^2) are relatively large. However, inspection of the modification indices reveals that it may be possible to improve Figure 1.

To explore the statistical consequences, some alternative models were constructed and checked against the model presented in Figure 1 and Table 5. An overview of the alternative models and their statistical outcomes is given in Table 6.

Table 6 – Comparison of different CFA models

Alternative measurement models	X^2	df	RMSEA	SRMR
1. Model with 4 latent factors (2 intrins., 2 extrins.; Figure 2)	1028.66	269	0.129	0.072
2a. As Model 1, but with error covariation Result_1 - Result_2	785.23	268	0.106	0.070
2b. Model with 5 latent factors (3 intrinsic, 2 extrinsic; Fig. 4)	758.39	265	0.104	0.077

In Table 6, model 1 is the model given in Figure 1 and Table 5. Model 2a of Table 6 allows correlation between result indicators Results_1 and Results_2. Compared to model 1, model 2a demonstrates a decrease in Chi-Square of 243.424 with a difference of only one degree of freedom (df). This difference between model 1 and model 2a is highly significant: model 2a results in a signif-

icant improvement in model 1. This is also shown in the values of RMSEA (0.106) and SRMR (0.070).

Additional explorative analysis of various parameters suggests combining intrinsic indicators Results_1, Results_2 and Discussion_2. This implies that there are three rather than two intrinsic latent factors, which changes the CFA model of Figure 1 into the CFA model of Figure 2 (see next page).

The statistical outcomes in Table 7 illustrate that, compared to CFA model 1, CFA model 2b (Figure 2) results in a significant improvement in Chi-Square (270.271; df=4; p<.01) and acceptable values for both RMSEA (0.104) and SRMR (0.077). Like the outcome of Table 5, this result concerning the relationships between intrinsic and extrinsic latent factors in Table 7 merits more attention for reasons of both interpretation and modelling in the EERQI conceptual framework. Moreover, this empirical outcome again supports the notion that using intrinsic indicators may add specific quality information to an EERQI consisting solely of extrinsic indicators and that introduction of extrinsic indicators may add specific quality information to an EERQI framework containing only intrinsic indicators.

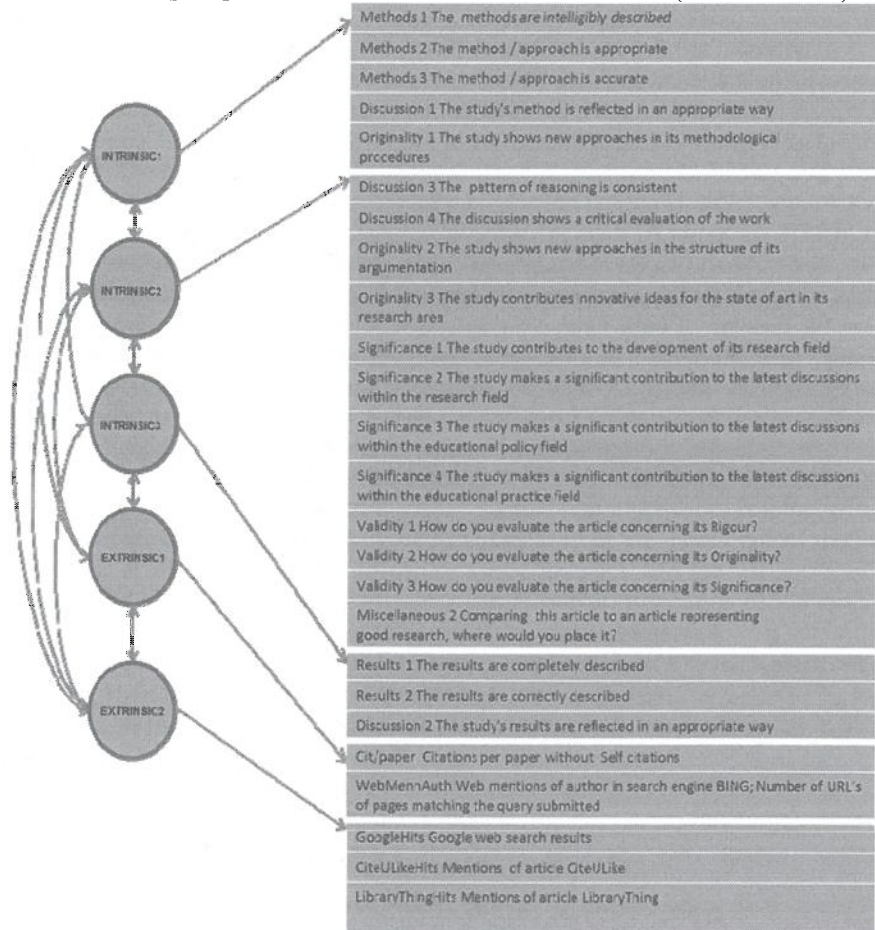## Figure 2 – Graphic presentation of CFA measurement model (5 latent factors)

Figure 2 – Graphic presentation of CFA measurement model (5 latent factors)

Table 7 – ML parameter estimates (standardised) of measurement model Fig. 2

	Factor loadings					
	INTRINS1: Methodol. adequacy	INTRINS2: Sign./orig./ consist.	IN-TRINS3: Results	EX-TRINS1: # cit./ WebBING	EXTRINS2: Hits 3 search	R2
Methods_1	0.907**					0.823*
Methods_2	0.862**					0.743*
Methods_3	0.914**					0.835*
Results_1			0.968**			0.937*
Results_2			0.975**			0.951*
Discussion_1	0.881**					0.776*
Discussion_2			0.787**			0.620*
Discussion_3		0.655**				0.429*
Discussion_4		0.611**				0.374*
Originality_1	0.787**					0.619*
Originality_2		0.796**				0.634*
Originality_3		0.873**				0.763*
Significance_1		0.900**				0.810*
Significance_2		0.911**				0.829*
Significance_3		0.809**				0.655*
Significance_4		0.721**				0.520*
Validity_1		0.542**				0.294*
Validity_2		0.785**				0.617*
Validity_3		0.842**				0.709*
Miscellaneous2		0.840**				0.705*
Cit/paper (sqrt)				0.591**		0.349*
WebMennAuth				0.686**		0.470*
GoogleHits					0.980**	0.960*
CiteULikeHits					0.674**	0.455*
LibraryTh-					0.803**	0.645*

	Factor covariances (correlations)			
	INTRINSIC1	INTRINSIC2	INTRINSIC3	EXTRINSIC1
INTRINSIC2	0.620**			
INTRINSIC3	0.740**	0.476**		
EXTRINSIC1	0.236	0.148	0.188	
EXTRINSIC2	0.146	0.085	0.113	0.460**

Fit indices: $X^2 (265)=758.385$ (p=0.000); RMSEA=0.104; SRMR= 0.077.
* $0.01 \leq p \leq 0.05$; ** $p<0.01$.

3.2. *Structural model of intrinsic and extrinsic latent factors*

A final exploration seeks to explain the latent factors within the CFA model in Figure 2. It is hypothesised that the degree to which the reviewed article or document is related to the reviewer's own area of research (item 21 or Miscellaneous1 in Table 1) influences the scores of the intrinsic latent factors. Inclusion of this explanatory variable in the CFA model of Figure 2 transforms this model into a causal or structural model. The causal relationships are represented by the three one-sided arrows between item 21 and the intrinsic latent factors: see the structural latent factor model in Figure 3.

Figure 3 – Structural model with intrinsic and extrinsic latent factors (indicators not included)
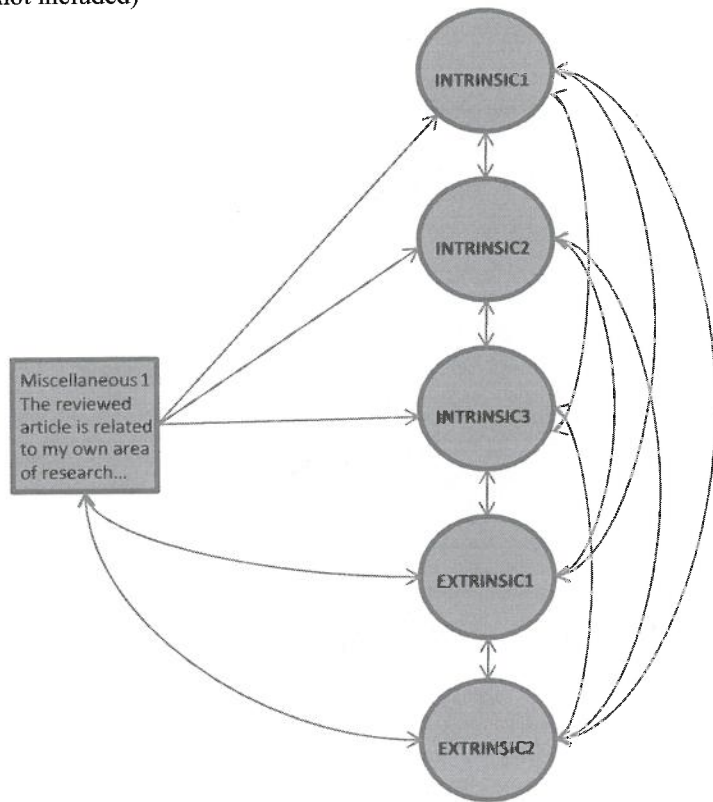
In Figure 3, the specific indicators for the latent factors are the same as those in Figure 2. Moreover, Figure 3 illustrates that the three intrinsic latent factors are regressed on the explanatory item Miscellaneous1 ('The reviewed article is related to my own area of research'). The correlations between the explanatory item and the two extrinsic factors are free to vary. The main results of Maximum Likelihood (ML) analysis using MPlus (version 6.1) are given in Table 8.

Table 8 - ML factor parameter estimates (standardised) of structural model

	Factor covariances (correlations)			
	INTRINSIC1	INTRINSIC2	INTRINSIC3	EXTRINSIC1
INTRINSIC2	0.604**			
INTRINSIC3	0.735**	0.463**		
EXTRINSIC1	0.247	0.162	0.195	
EXTRINSIC2	0.147	0.091	0.113	0.461**
	Direct effects				
	INTRINSIC1	INTRINSIC2	INTRINSIC3	EXTRINSIC1	EXTRINSIC2
Miscellaneous1	-0.176*	-0.247**	-0.128		
	Correlations				
Miscellaneous1				0.029	0.020

Fit indices: X^2 (284)=779.559 (p=0.000); RMSEA=0.101; SRMR= 0.075.
* $0.01 \leq p \leq 0.05$; ** $p < 0.01$.

Miscellaneous1 has significant effects on latent factors Intrinsic2 (-0.247) and Intrinsic1 (-0.176); the effect on Intrinsic3 (-0.128) is non-significant. This means that the more closely the reviewed document is related to the reviewer's own area of research, the higher the reviewer's evaluation scores with respect to significance, originality and consistency (Intrinsic2) and methodological adequacy (Intrinsic1). The two effects seem to reflect some subjective evaluation bias that may occur in peer reviewing.

Furthermore, Table 8 shows no significant statistical relationships between Miscellaneous1 and the extrinsic factors. The correlation between intrinsic factors – and not extrinsic factors – and the relevance of the reviewed document to the reviewer's own area of research supports the validity of the conceptual framework.

4 Conclusions

4.1 An initial prototype framework of EERQI indicators

A pilot was conducted to collect data with respect to 'intrinsic' and 'extrinsic' indicators of research documents. The research question was: What do statistical analyses show us about the relationships between the two types of indicator and what does this mean when constructing a prototype EERQI framework? To answer this question, peer review scores or intrinsic indicators were used to assess aspects of research articles or documents related to methodology, results, discussion, originality, significance, validity, and miscellaneous matters. Extrinsic indicators for the same documents were Web-based.

Some consecutive measurement models and their empirical results confirmed the potential relevance and functionality of intrinsic latent factors, extrinsic latent factors, and their indicators. A final check was whether the degree to which a reviewed article is related to the reviewer's own area of research influences the scores of the intrinsic latent factors. Empirical testing in a causal structural model indeed revealed that the more the reviewed document is related to the reviewer's own area of research, the higher reviewer's evaluation scores with respect to 1) significance, originality and consistency, and 2) methodological adequacy. There are no significant relations between the reviewer's own area of research and the extrinsic factors.

The differentiated relationships and outcomes of this pilot support the validity of both the conceptual framework and the empirical research. The conclusion is that an initial prototype EERQI framework has been constructed. The relevant conceptual framework is presented in Figure 3 and Figure 2 combined. It is possible that other types of information, for example semantic indicators and factors, can be integrated into these figures and follow-up research.

Given the statistical outcomes related to Figures 3 and 2 in Tables 5 – 8, a further conclusion is that an initial empirical test of the conceptual EERQI framework has been successful. The main goal of the EERQI project – to improve citation-only assessments of the quality or impact of educational and other research – has been supported. It is, however, important to describe some limitations of the study.

4.2 Methodological limitations

From a methodological point of view, the present pilot has a number of limitations which any follow-up analyses or research should seek to eliminate. These include:

- the exact operationalisation and assessment of both extrinsic and intrinsic indicators need careful consideration, for example for reasons of validity and representativeness;
- the pilot covered only a fairly small number of documents and reviewers;
- the ratio of number of documents to number of indicators (171:25) is relatively low;
- the distribution characteristics of the variables and their consequences for statistical analysis merit greater attention;
- the interobserver reliability of the reviewing information also merits greater attention;
- possible effects of language differences have not been taken into account;
- multilevel analysis was not applied because of the small number of documents/reviewers.

4.3 Future steps

The EERQI project has many different sides to it and considerable potential. In the future, more of the project partners and other parties may be convinced of its significance, originality and consistency (latent factor Intrinsic 2). Exploiting its potential and improving the focus on both methodological adequacy (latent factor Intrinsic 1) and semantic indicators and latent factors should optimise the steps to the further construction and use of EERQI.

5 References

Bridges, D. (2009): Research quality assessment: impossible science, possible art? British Educational Research Journal, 35(4), 497-517.

EERQI project (2010): State of the art reports on EERQI project parts. Preparatory meeting for the second EERQI workshop, 18-19 March. EERQI Project Report.

Gogolin, I. (2008): European Educational Research Quality Indicators (EERQI). (Project 217549). FP7 Collaborative project. Hamburg, Universität Hamburg.

Gradmann, S., Sieber, J., & Stoye, D. (2011): Extrinsic indicators used in EERQI. Berlin: Humboldt-Universität, Institut für Bibliotheks- und Informationswissenschaft.

Mooij, T. (2008a): Suggestions for a first conceptual framework to construct EERQI. Contribution to the international project 'European Educational Research Quality Indicators' (Project 217549; FP7). Nijmegen, The Netherlands: Radboud University, ITS.

Mooij, T. (2008b): Intermediate conceptual framework and procedures to construct EERQI. Contribution to the international project 'European Educational Research Quality Indicators' (Project 217549; FP7). Nijmegen, The Netherlands: Radboud University, ITS.

Mooij, T. (2011): European Educational Research Quality Indicators (EERQI): A first prototype framework of intrinsic and extrinsic indicators. Paper for the final conference in Brussels, University Foundation, 15-16 March 2011 of the international collaborative project 'European Educational Research Quality Indicators' (EU Project 217549; FP7). Nijmegen, The Netherlands: Radboud University, ITS.

Nolin, J., & Åström, F. (2010): Turning weakness into strength: Strategies for future LIS. Journal of Documentation, 66(1), 7-27.

Sándor, Á., & Vorndran, A. (2009): Detecting key sentences for automatic assistance in peer reviewing research articles in educational sciences. EERQI Project Report.

Sieber, J., & Stoye, D. (2011): Description of a Measure. Berlin: Humboldt-Universität, Institut für Bibliotheks- und Informationswissenschaft.